Representing Classes of Things and Properties in General in Conceptual Modelling: An Empirical Evaluation

Graeme Shanks, University of Melbourne, Australia

Daniel Moody, University of Twente, The Netherlands

Jasmina Nuredini, Simsion and Associates, Australia

Daniel Tobin, University of Melbourne, Australia

Ron Weber, Monash University, Australia

ABSTRACT

How classes of things and properties in general should be represented in conceptual models is a fundamental issue. For example, proponents of object-role modelling argue that no distinction should be made between the two constructs, whereas proponents of entity-relationship modelling argue the distinction is important but provide ambiguous guidelines about how the distinction should be made. In this paper, the authors use ontological theory and cognition theory to provide guidelines about how classification should be represented in conceptual models. The authors experimented to test whether clearly distinguishing between classes of things and properties in general enabled users of conceptual models to better understand a domain. They describe a cognitive processing study that examined whether clearly distinguishing between classes of things and properties in general impacts the cognitive behaviours of the users. The results support the use of ontologically sound representations of classes of things and properties in conceptual modelling.

Keywords: *Conceptual Modelling, Entity, Information Systems Development, Object-Role Model, Ontology, Normalization, Property*

1. INTRODUCTION

The notions of classes of things and the properties that things in the class possess (properties in general) have been of interest to philosophers concerned with ontology (the nature of the world) (e.g., Bunge, 1977). They have also been of interest to information systems researchers and practitioners concerned with finding better ways to model the world. For instance, the representation of classes of things and properties in general features in early work on conceptual

DOI: 10.4018/jdm.2010040101

modelling (Chen, 1976; Nijssen, 1976; Kent, 1978). It also features in more-recent object-oriented conceptual modelling approaches–in particular, the Unified Modelling Language (e.g., Rumbaugh et al., 1999).

For a number of reasons, the notions of classes of things and properties in general and their representation in conceptual models are problematic. First, not all scholars agree that things and properties are distinct phenomena. For instance, nominalist philosophers "dispense with properties, which they regard as Platonic fictions, and attempt to reduce everything to things, their names, and collections of such" (Bunge, 1977, p. 57). Moreover, those philosophers who do sustain a distinction between things and properties face the difficult task of showing how the distinction should be made (e.g., Denkel, 1996).

Second, some information systems scholars argue the distinction between classes of things and properties in general ought not to be sustained in conceptual models, because different users may perceive the same phenomena differently (in short, implicitly these scholars subscribe to a nominalist philosophy). For example, in the object-role approach to conceptual modelling, the distinction between classes of things and properties in general is not maintained (Halpin, 2008). Both are represented using the object symbol in a conceptual model. Similarly, Date (2003, p. 436) eschews the distinction between an entity (thing) and a relationship (type of property of a thing): "In this writer's opinion, any approach that insists on making such a distinction is seriously flawed, because…*the very same object* can quite legitimately be regarded as an entity by some users and a relationship by others."

Third, even when conceptual modelling approaches allow classes of things to be distinguished from properties in general, how the distinction should be maintained is often unclear. In the entity-relationship (ER) model (Chen, 1976), for example, classes of things are supposed to be represented as entity types, and properties in general are supposed to be represented as attribute types. Nonetheless, entity-type symbols are often used to represent both classes of things and properties in general. For instance, a *preference*, which many individuals would deem to be a property in general of a class of things, might be represented as an entity type that is connected via a relationship type to a *client* entity type (see, e.g., Connolly & Begg, 2005, p. 344).

Fourth, disputes arise about how classes of things and properties in general should be represented in conceptual models if database design considerations are to be taken into account. For example, Simsion and Witt (2001, p. 104) state: "Attributes in an ER model correspond to columns in a relational model." They further suggest that ER models should be "normalized" and repeating groups of attributes removed to form additional entity types. Thus, they argue that representations in a conceptual model ought to be influenced by database design considerations.

A conceptual model is used to discover and document user views of an information system and to provide a basis for informed discernment, reconciliation, and compromise among users and information systems professionals (Hirschheim et al., 1995). Therefore, we argue that the representation of classes of things and properties in general in conceptual models should be based on a sound underlying theory about the structure and dynamics of phenomena in the world (Parsons & Wand, 2008). In this regard, however, little empirical work has been done (Evermann & Wand, 2006; Moody, 2002; Weber, 1996). Consequently, we undertook a theoretically based, empirical evaluation of alternative conceptual-modelling representations of classes of things and properties in general.

Five factors motivated our work. First, it is well recognized that the cost of fixing errors grows exponentially the later they are discovered in the system development process (e.g., Boehm, 1981). Because conceptual modelling work is undertaken early in the system development process, improvements in conceptual modelling practice potentially should lead to high payoffs (Moody & Shanks, 1998).

Second, in the context of implementing, operating, and maintaining enterprise systems, conceptual models are becoming increasingly important. They provide a means of evaluating the "fit" between an organization's needs and the business models embedded within the enterprise application software used to implement such systems (Sia & Soh, 2002). Similarly, in the context of implementing, operating, and maintaining interorganizational information systems, conceptual models provide a means to compare and contrast the different business models that underlie the participants' operations.

Third, we sought to test previous theoretical work undertaken to predict how well different types of representations facilitate or inhibit human understanding of real-world phenomena (e.g., Weber, 1997). If we can make accurate predictions about what types of conceptual modelling practices are likely to work well, we avoid the high costs associated with learning about the strengths and weaknesses of different practices through experience.

Fourth, it is important to determine which type of representation of real-world phenomena in a conceptual model enables humans to better understand the phenomena and why this outcome occurs. When conceptual models are prepared initially (e.g., by systems analysts), users of an information system are asked to evaluate the models to determine how accurately and completely the models represent their perceptual worlds. If users cannot understand a conceptual model clearly in the first place, their ability to validate the model is impaired. Moreover, subsequent users may employ conceptual models to understand the functionality provided by an information system. If these users cannot understand the conceptual models clearly, their ability to engage effectively with the information system is undermined.

Fifth, we sought to contribute to improved conceptual modelling practice. As we discussed above, many different, sometimes-ambiguous guidelines for representing classes of things and properties in general appear in the practitioner literature. For example, the object role model does not distinguish between classes of things and properties in general (Halpin, 2008), whereas proponents of the entity relationship model provide differing guidelines about how the distinction is made (Chen, 1976; Simsion & Witt, 2001). These may confuse rather than assist practitioners (Simsion & Witt, 2001). We have designed our study to enable a comparison of alternative representations of classes of things and properties in general to determine empirically which representation supports better user understanding. If we develop improved conceptual-modelling rules for classes of things and properties in general, the conceptual-modelling tasks that practitioners undertake should be more straightforward.

The remainder of our paper proceeds as follows. The next section discusses the theory and proposition that underpin our empirical work. The third and fourth sections describe the design, conduct, and results of the laboratory experiment and cognitive process tracing study we undertook. The fifth section discusses our empirical results and relates the results of the process tracing study to those of the experiment. The sixth section presents our views on the implications of our results for research and practice. Finally, we discuss some limitations of our research and some directions for future research.

2. THEORY AND PROPOSITION

We base our proposition on two theories. First, we use the ontological theory proposed by Bunge (1977) and adapted for information systems by Wand and Weber (1993) to argue that a distinction between the representation of classes of things and properties in general is essential to avoid construct overload. We also use a theory of cognitive clustering within human information processing (Bousfield, 1953; Miller, 1956; Baddeley, 1994) to argue that clustering properties in general with the classes of things to which they belong helps humans understand complex representations.

The ontological theory we use analyzes the representation of classes of things and properties in general as follows:

1. "The world is made of things that possess properties" (Wand et al., 1999, p. 497). These are the two basic constructs that are needed to describe the world. There can be no bare things; they must possess one or more properties. Properties cannot exist by themselves; they must be attached to one or more things. Furthermore, properties themselves may not have properties (Wand et al., 1999, p. 498).
2. There are two types of properties: intrinsic properties and mutual properties. Intrinsic properties depend on one thing only–for example, the height of a person (Wand et al., 1999, p. 498). Mutual (or relational) properties depend on two or more things–for example, being a university student depends on both a person and a tertiary institution (Wand et al., 1999, p. 498).
3. Things can interact with each other (Wand et al., 1999, p. 503). Two things interact (are coupled) when a history of one thing (manifested as a sequence of the thing's states) would be different if the other thing did not exist. The existence of a mutual property between two things can indicate that they interact with each other. Mutual properties that manifest interactions between two things are called *binding mutual properties.* For example, the mutual property that a person is a student at a university implies that the existence of the university affects the state of the person (and vice versa). If the university ceases to exist, the state of the person changes from being a student. If the person leaves the university, then the state of the university changes (the list of students will change in value (adapted from Wand et al., 1999, p. 503).
4. Properties (represented by attributes) that belong to the substantial individuals of all members of a set, *S*, are *properties in general* (e.g., age); properties that belong to a specific individual in the set are *properties in particular* (e.g., age is 16 years) (Bunge 1977, p. 63).

In the context of ontological theory we use, a property in general should not be represented as an entity type (or class) in conceptual modelling[1]. This practice leads to construct (semantic) overload because the same grammatical construct (an entity type or class symbol) has been used to represent two ontological constructs (classes of things and properties in general). Under these circumstances, users of the model must employ tacit knowledge to determine the semantics of the model (Wand & Weber, 1993).

It is well known that humans cognitively cluster phenomena that they perceive to be related in some way (Bousfield, 1953; Miller, 1956; Baddeley, 1994). They appear to use clustering as a means of dealing with the complexity they often encounter in their perceptual worlds. By focusing on the cluster rather than each phenomenon that makes up the cluster, they reduce cognitive load and enhance their abilities to make sense of the world. Parsons and Wand (2008a, 2008b) highlight the importance of classification as a cognitive clustering mechanism within science generally.

By sustaining a distinction between classes of things and properties in general, we argue humans invoke a cognitive strategy that allows them to deal with complexity. Properties in general "naturally" cluster with the things in the class to which they belong. Thus, perceiving the world in terms of classes of things and their properties in general helps humans to mitigate the cognitive problems they experience when they perceive phenomena to be complex.

For example, in the context of conceptual modelling, Moody (2002) and Weber (2003) argue information systems analysts and users who have to undertake the often-difficult task of decomposing an application domain into systems and subsystems will achieve a better outcome if the conceptual model developed to represent the domain clearly distinguishes between classes of things and properties in general in the domain. Similarly, Weber (1996) found that students trained in object-role modelling, which does not distinguish between classes of things and their properties in general, still used clusters of things and their properties as a basis for recalling object-role models (obtained from two organizations) that they had studied. He proposes a model based on spreading activation theory (Collins & Quillian, 1969; Anderson, 1983; Anderson & Pirolli, 1984) to account for why clustering of things with their properties will facilitate comprehension and recall of conceptual models.

Although the ontological theory we use provides a means of distinguishing and representing classes of things and properties in general in conceptual modelling, some conceptual modelling approaches do not sustain the distinction–for example object-role modelling (Halpin, 2008). For this reason, we argue that it is important to evaluate empirically the consequences of sustaining or not sustaining a distinction between classes of things and properties in general. Although theories of representation can never be proved correct, we can test falsifiable propositions based upon them (Bacharach, 1989; Popper, 1961).

In light of the ontological theory and the theory of cognitive clustering we use, we contend that conceptual models ought to maintain a distinction between classes of things and properties of things because it will allow their users to better comprehend the perceptual worlds that the models are supposed to represent. Thus, the following proposition motivates the empirical work we undertook:

Proposition: Conceptual models that distinguish between classes of things and properties in general will enable their users to better understand the semantics of the perceptual domains the models are representing than conceptual models that do not sustain this distinction.

3. LABORATORY EXPERIMENT

We employed a laboratory experiment because we sought to (a) manipulate in specific ways those phenomena about a domain that we represented in a conceptual model to try to obtain support for a cause-effect relationship, (b) control for extraneous factors that might confound any impacts of alternative representations of classes of things and properties in general in conceptual models on how well users of the models understood these constructs, and (c) obtain sufficient numbers of participants in our research to test statistically hypotheses motivated by our proposition.

3.1 Design and Measures

A four-group, post-test only experimental design was used with one active between-groups factor. This factor, "type of representation," had four levels. The first, which we term the "ontologically sound" level, represented classes of things as entity types and properties in general as attribute types in an ER diagram. The second, which we term the "partially ontologically sound" level, represented only *mutual* properties in general (properties of *n*-tuples of classes of things) as entity types. *Intrinsic* properties in general (properties inherent to single class of things) were still represented as attribute types. The third, which we term the "normalized" level, represented mutual properties in general and some intrinsic properties in general as entity types. This level complied with the approach to representing application domains via ER diagrams used by many practitioners (Simsion & Witt, 2001). The fourth, which we term the "entity-only" level, represented both classes of things and properties in general as entity types. This level follows the principles used by object-role modellers (Halpin, 2008).

The dependent variable, performance, was evaluated in three ways: comprehension performance, problem-solving performance, and discrepancy-checking performance. These are all measures of script interpretation in the evaluation framework proposed by Gemino and

Wand (2004). *Comprehension* involves someone using a conceptual model to understand the "surface-level" features of a domain. *Problem solving* involves someone using a conceptual model to solve problems that might arise in the domain. We differ from earlier work (e.g., Bodart et al., 2001; Gemino, 1999; Mayer, 1989) in that our problem-solving questions involved using and navigating a conceptual model to understand more-complex aspects of a domain, thereby providing a better indicator than comprehension of someone's "deep" understanding of a conceptual model. We used comprehension and problem-solving tasks to test how well the four conceptual models communicated the semantics of a domain to the participants in our experiment. *Discrepancy checking* involves someone comparing a conceptual model against a textual description of the domain (into which differences or discrepancies have been added) to evaluate whether the conceptual model represents the semantics manifested in the textual description accurately and completely (Moody, 2002). This task provides an alternative to answering questions as a means of testing someone's understanding of a conceptual model.

We measured comprehension, problem-solving, and discrepancy-checking performance in two ways: (a) accuracy, and (b) time taken. *Comprehension accuracy* was measured via the number of comprehension questions answered correctly by a participant. *Problem-solving accuracy* was measured via a score that was based on whether participants obtained a correct answer to the problem and provided a clear explanation of their rationale. *Discrepancy-checking accuracy* was measured via a score that was based on whether participants (a) identified correctly a discrepancy between the conceptual model and some text that described part of the application domain represented by the conceptual model, and (b) provided a clear explanation of the nature of the discrepancy. *Comprehension time*, *problem-solving time*, and *discrepancy-checking time* were all measured via the number of minutes (or part thereof) that participants took to complete each task.

3.2 Materials

Seven sets of materials were developed for the experiment. We present them below in three sub-sections: profile and training materials, conceptual models, and understanding tasks materials.

3.2.1 Profile and Training Materials

Two sets of profile and training materials were developed. The first comprised a "personal-profile" questionnaire to obtain information about participants' academic qualifications, industry experience, work experience, time in their current position, and modelling experience. We used these materials to ensure participants who received the different treatments had similar academic qualifications, work experience, etc.

The second was a summary of the ER symbols used in the diagrams provided to participants in the experiment. Note, to maximize our contribution to conceptual modelling practice, we decided to base our study on the ER approach to conceptual modelling. This approach is used widely in practice (Rosemann et al., 2003; Simsion & Witt, 2001).

3.2.2 Conceptual Models

Four ER diagrams of alternative conceptual models of a sales order domain (one that is understood widely) were developed. We first prepared a diagram using an approach that is employed widely in practice (Figure 1)–namely, where entity types essentially are third normal form relations (Simsion & Witt, 2001). Using this approach, classes of things are represented as entity types. In addition, multi-valued attributes in general (intrinsic properties in general) are represented as entity types (known as attributive or characteristic entity types–for example, *Customer Contact Person* in Figure 1). Similarly, value domains are also represented as entity types (known as classification entity types–for example, *Customer Industry Type* in Figure 1), and many-to-many relationships (mutual properties in general) are represented as entity types (known as intersection or associative entity types–for example, *Sales Order Item* in Figure 1). In ontological terms, many ontological constructs are represented by one modelling construct, an entity type, which leads to construct overload (Wand & Weber, 1993). In developing the model in Figure 1, we first analyzed a typical model from practice to work out the ratios of the different categories of entity types described above. We ensured our model had similar ratios to increase its external validity.

Next, we developed an ontologically sound model of the sales order domain. In preparing the ER diagram for this model, we adopted a two-stage approach. First, we transformed the "normalized" ER model into a "partially ontologically sound" model (Figure 2) by removing the attributive and classification entity types and folding their attributes in general into the related entity type (e.g., attributes in general from *Customer Contact Person* and *Customer Industry Type* were folded into the *Customer* entity type). These transformations are consistent with ontological principles for representing intrinsic properties in general.

Third, we transformed the "partially ontologically sound" model into the "ontologically sound" ER model (Figure 3) by removing the associative entity types through folding their attributes in general into both related entity types (e.g., attributes in general from *Sales Order Item* were folded into both the *Order* entity type and the *Product* entity type). This transformation is consistent with ontological principles for representing mutual properties in general[2]. When these transformations are made, minor information losses occur that are associated with constraints on relationships that were deleted. We were careful to avoid involving these aspects of the models in our comprehension, problem-solving, and discrepancy-checking tasks to ensure information equivalence issues did not confound the results (Burton-Jones et al., 2009; Parsons & Cole, 2005; Siau, 2004).

We then developed a model of the sales domain that does not distinguish between things and properties (Figure 4). We transformed the "normalized" ER model by creating a new en-

Figure 1. Normalized ER model

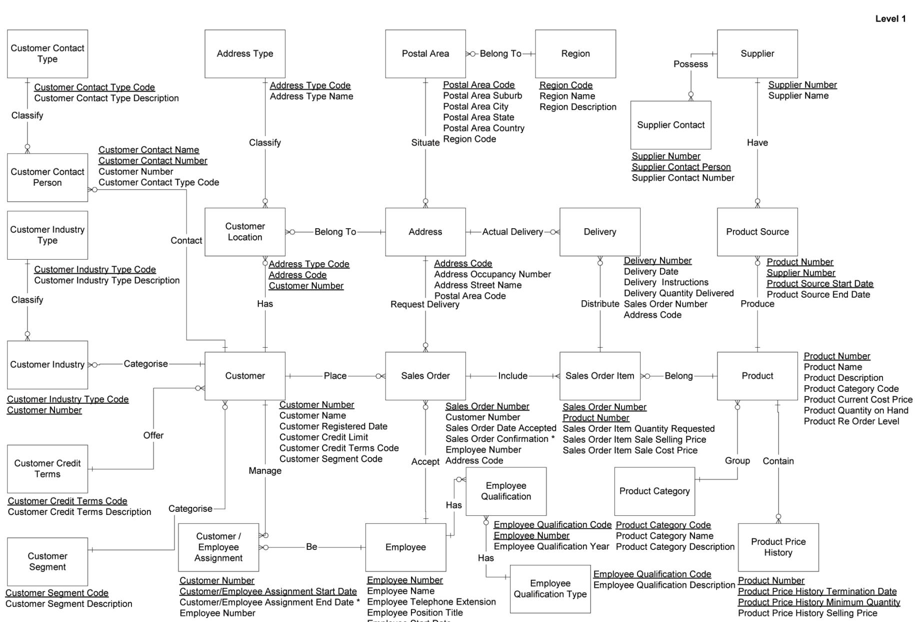


Figure 2. Partially ontologically sound ER model

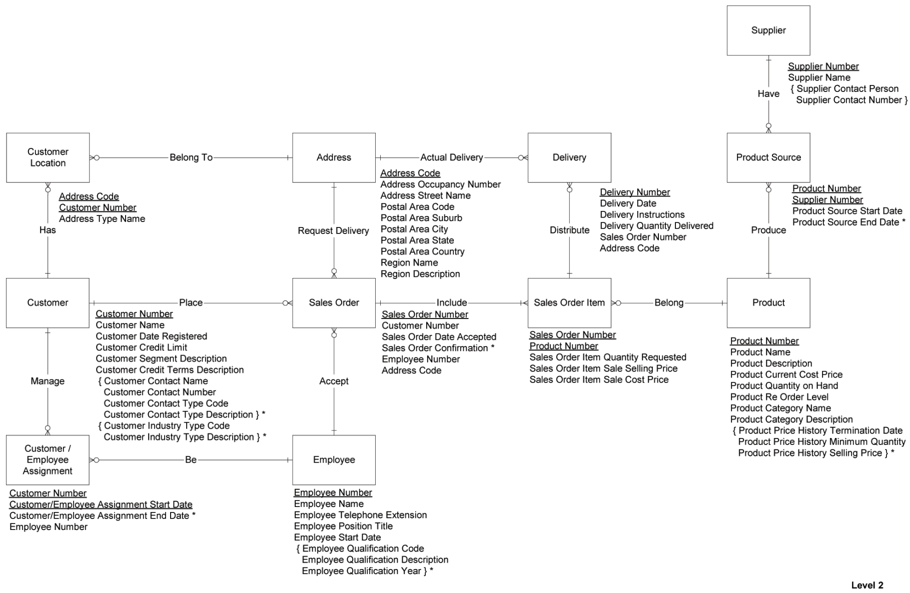
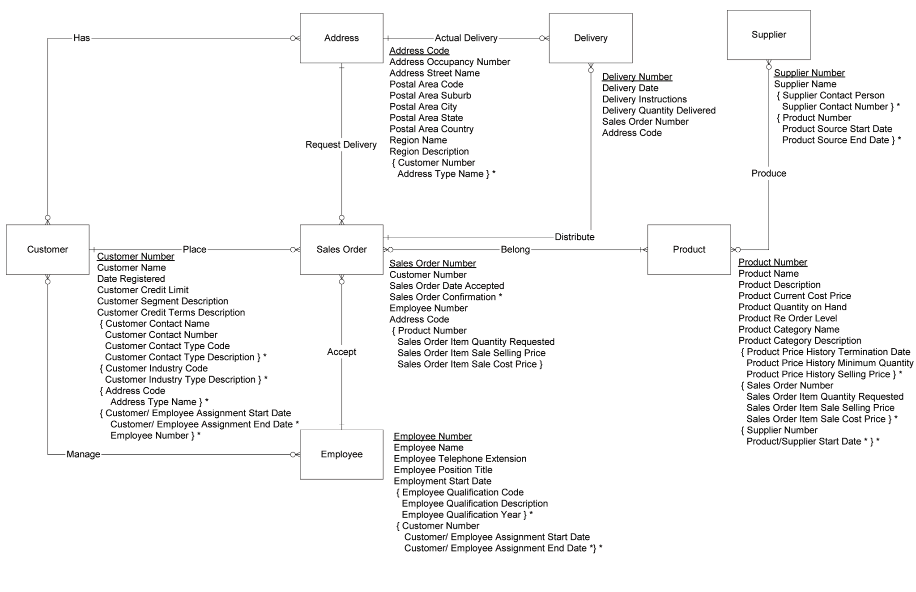
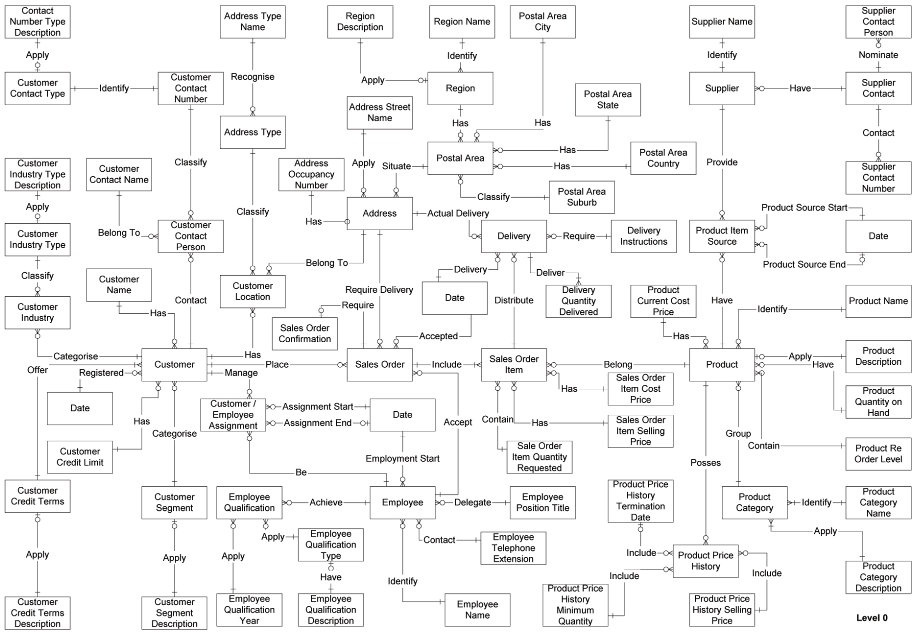


Figure 3. Ontologically sound ER model



tity type for each attribute. This transformation is consistent with the philosophy underlying object-role modelling–namely, that no distinction should be made between classes of things and their properties in general. "Facts" that connect classes of things are the key concept.

Figure 4. Entity-Only ER model



When this transformation was made, a more complex model resulted. Nonetheless, the cardinality constraints on the relationships provided clear semantics.

The four categories of model used in this study constitute four points in a continuum (Figure 5) varying from the "entity-only" ER model, where no distinction is made between classes of things and properties in general, to the "normalized" ER model, where all mutual properties in general but only some intrinsic properties in general are represented as entity types, to the "partially ontologically sound" ER model, where mutual properties in general are represented as entity types, through to the "ontologically sound" ER Model, where a clear distinction is made between classes of things and properties in general[3]. Table 1 shows the

Figure 5. Thing/Property continuum

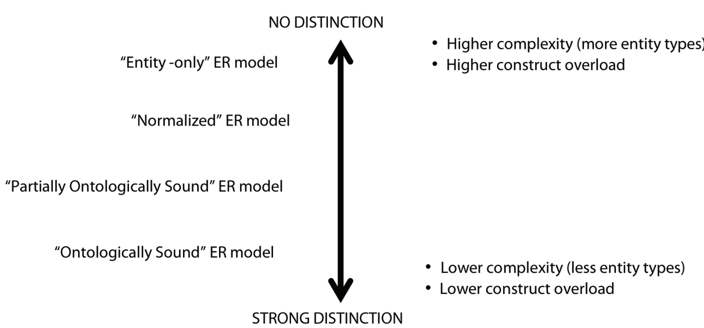


Table 1. Mapping between models

Ontological Concept	Ontologically Sound ER Model	Partially Ontologically Sound ER Model	Normalized ER Model	Entity-Only ER Model
Thing	Entity	Entity	Entity	Entity
Intrinsic Property	Attribute	Attribute	Entity or Attribute	Entity
Mutual Property	Attribute	Entity or Relationship	Entity or Relation-ship	Entity
Value Domain	Domain	Domain	Domain	Entity

mapping from ontological concepts to modelling notation constructs for each of the four types of model.

In preparing these models, a problem we faced was how we should distinguish classes of things from properties in general in our application domain. Clearly, if this task were straightforward, no debate would arise about whether classes of things and properties in general were distinct phenomena in the world. To make the distinction, we used three criteria. First, qualitatively we deem a thing to be a physically independent phenomenon in the world–that is, it satisfies the condition that it is "capable of existing in physical space, by itself, without requiring the support of anything else" (Denkel, 1996, p. 16). For example, a particular person is a thing because the person is a phenomenon that exists independently in physical space. A skill, however, is a property because it cannot exist independently in physical space–it must inhere in a thing. In this regard, if we move a person in physical space, the fact that the person's skills also move is "gratuitous." If we want to move the person's skills in physical space, however, we have to move the person "in order to" accomplish this outcome.[4]

Second, if the phenomenon when named can be conceived as a function that maps something to a value domain, we deem it to be a property in general. For instance, "person" cannot be conceived as a function that maps something to a value domain. On the other hand, skill is a property in general of the "person" class of things because it can be conceived as a function that maps something to a value domain (e.g., skill can take on the values "programming," "accounting," and "bass playing" for a particular person thing). In short, we cannot assign a thing a value; only its properties can be assigned values.

Third, following Denkel (1996, p. 35), things survive a change of their properties,[5] whereas properties cannot survive changes of things. For instance, a person may have the skill of being a bass player. The skill may be lost, however, if the person fails to practice the instrument for a long period or suffers permanent injury to a hand. Nonetheless, the person survives in spite of the skill being lost. The person's skill of being a bass player cannot simply be transported or given to another person. Note, we are focusing on a *property in particular* (a particular person's particular skill as a bass player) as opposed to a *property in general* (the skill of bass playing that many people possess) (Bunge, 1977, pp. 62-65).

3.3.2 Understanding Tasks Materials

Three sets of materials were developed. The first comprised 10 comprehension questions. They were designed to test a user's ability to access and navigate the model for relatively simple, surface-level tasks. Responses to questions were "yes," "no," or "not sure" (included to minimize guessing). An example is: "Can an employee be assigned to manage more than one customer at a time?" (See Appendix 1 for more examples.)

The second comprised 10 problem-solving questions. They were designed to force participants to use the ER diagrams in deeper ways and to obtain a correct answer based on the diagrams rather than relying on their tacit knowledge of the sales order domain. Responses to questions were "possible," "not possible," or "not sure" (included to minimize guessing). Participants also had to provide a brief explanation of their answer. An example is: "An Ontological Plastics supplier wishes to send samples of new and improved hoses to customers who regularly order hoses. Can we determine the number of hoses each customer has had delivered in the previous three months and the date of each delivery?" (See Appendix 1 for more examples.)

The third comprised a typed transcript of a fictitious interview between a conceptual modeller and three users. The users described a number of aspects of the sales order domain. For example, a comment made by the second user is: "Yes, well, each of our deliveries consists of a single product but may contain items from multiple sales orders, so we need to know precisely where they all need to go." Beside the transcript was a column where participants could note any discrepancies they identified between the semantics in the conceptual model they had been given and the details of the sales domain described in the transcript. We had seeded the transcript so its semantics differed in eight ways from the conceptual model. Again, participants had to provide a brief explanation of any discrepancy they identified. (See Appendix 1 for an example paragraph from the transcript.)

3.3 Participants

Participants in the experiment were 80 volunteers who either worked in industry or had worked in industry but at the time of the experiment were postgraduate students. None performed or had performed information systems/information technology functions as their primary role within their organization. In essence, in the experiment they acted as surrogate end users. With the exception of one person who had 20 years' work experience, all had at least an undergraduate degree with majors in diverse areas (e.g., arts, architecture, psychology, law, accounting, education, mathematics, information systems, engineering). Forty-two had between one and five years' work experience, and 14 had in excess of 10 years' work experience. Sixty-one had no experience of data models. The remainder had minor experience of one or two modelling techniques like flowcharts or financial models. Each participant was paid $30 to undertake the experiment.

3.4 Procedures

Participants were first assigned randomly to one of the four treatments (20 per treatment). Sixty-nine participants were run singly through the experiment, seven undertook the experiment together in a group, and there were two other groups of two participants[6]. When they arrived to undertake the experiment, they were asked to complete a consent form and the personal-profile questionnaire.

Next they were given the document that explained the ER symbols. Participants were permitted to discuss the symbols with the researchers until they indicated they felt confident in their understanding of the ER symbols. They retained and could refer to the ER summary throughout the experiment.

When participants indicated they were ready to begin, they were given the "ontologically sound" ER diagram, the "partially ontologically sound" ER diagram, the "normalized" ER diagram, or the "entity-only" ER diagram, depending on the treatment to which they had been assigned randomly. They retained and could refer to the diagram throughout the experiment. The times taken to answer each comprehension and problem-solving question were recorded as well as the total time taken to perform the discrepancy-checking task.[7] Notes were also made based on participant reactions, queries, and approaches to each question. With 67 participants, one researcher conducted the experiment, while another took notes, recorded times, and observed the participant's behavior during the experiment. The remaining 13 participants recorded their own times because they undertook the experiment in groups or only one researcher could be present as two participants were undertaking the experiment concurrently in different locations. Overall, the experiment took about 90 minutes to complete.

Note that we did not randomize the order in which participants were given the comprehension, problem-solving, and discrepancy-checking tasks (see, also, Mayer, 1989; Mayer & Gallini, 1990). Rather, we had participants follow a sequence of tasks aimed at testing the different types of understanding we expected they would acquire at different stages as they progressively came to grips with the meaning of the conceptual model they had been given. In this regard, at the outset we expected that participants would first acquire a surface-level understanding of the conceptual model. Hence, we gave them the comprehension task first to test how well they had acquired a surface-level understanding of the domain using the four models. Next, we expected that participants would build on their surface-level understanding to develop a deep-level understanding of the conceptual model. Ideally, we would have split our participant group to then undertake either the problem-solving or the discrepancy-checking task. We lacked the resources to pursue this strategy, however. Thus, we chose to give participants the problem-solving task before the discrepancy-checking task because we thought the former would provide us with a more-valid and more-reliable measure of participants' deep-level understanding of the conceptual model.

3.5 Results

Scores for the individual items on the comprehension, problem-solving, and discrepancy-checking dependent measures were first calculated. Next, a reliability check on the dependent measures was undertaken. Statistical analyses were then performed on the scores for each dependent measure to test the proposition that underlies our research.

3.5.1 Data Scoring

Scores were awarded as follows (and then normalized to a score out of 100):

1. *Comprehension* (10 questions; maximum score 10)

One mark was given if the answer ("possible" or "not possible") was correct; zero was given if a participant selected "not sure" or their answer was incorrect. Participants were encouraged to answer "not sure" rather than guess an answer.

2. *Problem Solving* (10 questions; maximum score 20)

Two marks were given if the answer ("possible" or "not possible") was correct; zero was given if a participant selected "not sure" or their answer was incorrect. Explanations were used to amend the score only if the explanation was inconsistent with the answer given. If the answer was correct but the explanation was unclear and did not support the answer, one mark was subtracted from the score. If the answer was incorrect or "not sure" but the explanation indicated the participant was reasoning coherently about the problem, one mark was added to the score. The two of us who conducted the experiment independently scored the problem-solving measures on pre-formatted scoring sheets.[8] Few differences arose between the two sets of scores. Where they did occur, they were discussed and reconciled.

3. *Discrepancy Checking* (8 discrepancies; maximum score 16)

One mark was given if a participant correctly identified a discrepancy between the text and the conceptual model. A second mark was given if the participant then provided a clear explanation of the nature of the discrepancy. Again, the two of us who conducted the experiment independently scored the discrepancy-checking measure. Few differences arose, and they were discussed and reconciled where they did occur.

3.5.2 Reliability Analysis of Dependent Measures

Cronbach alphas for the comprehension, problem-solving, and discrepancy-checking measures were .56, .42, and .56. Given the complex, multifaceted "understanding" construct that underlies these measures, we believe their reliability is satisfactory (Nunnally, 1978). Deletion of any "item" from the measures neither increased nor decreased alpha markedly.

3.5.3 Tests of Proposition

In this section, we report the results of the tests we undertook of the proposition that "Conceptual models that distinguish between classes of things and properties in general will enable their users to better understand the semantics of the perceptual domains the models are representing than conceptual models that do not sustain this distinction."

For each of the four models (ontologically sound, partially ontologically sound, normalized, and entity-only), we test the accuracy (interpretational fidelity) and time taken (interpretational efficiency) (Burton-Jones et al., 2009) for the comprehension, problem solving, and discrepancy checking tasks. We first report the accuracy and time measures and their correlations. We then explain the statistical analyses we undertook. Finally, we present our significant findings.

1. Accuracy and Time Measures

Table 2 shows the means and standard deviations for the accuracy and time measures associated with the three primary performance constructs (comprehension, problem solving, and discrepancy checking). Table 3 shows the Pearson correlation coefficients among the accuracy and time measures.

2. Statistical Analyses

The three accuracy measures and the three time measures are moderately correlated with one another (see Table 3). For this reason, we undertook two separate, single-factor multivariate analyses of variance (MANOVA). In both MANOVAs, the factor was type of model at four levels (ontologically sound, partially ontologically sound, normalized, and entity-only). In the first, the dependent measures were comprehension accuracy, problem-solving accuracy, and discrepancy-checking accuracy. In the second, the dependent measures were comprehension time, problem-solving time, and discrepancy-checking time. Checks of the assumptions underlying both MANOVAs (univariate and multivariate normality, univariate and multivariate outliers, linearity, homogeneity of variance-covariance matrices) revealed no violations.

For the three *accuracy* measures, the model was significant: $F(9, 180.247) = 2.748$, $p = .005$; Wilks' Lambda = . 980, partial eta squared = .099. In this light, we used the Roy-Bargmann stepdown analysis procedure to determine which of the three accuracy measures were statistically significant (Tabachnick & Fidell, 2007, pp. 271-272). We entered the dependent variables into the stepdown analysis following the order in which they had been measured in the experiment (i.e., comprehension accuracy, problem-solving accuracy, and discrepancy-checking accuracy).

We first undertook a univariate analysis of variance (ANOVA) with type of model as the factor and comprehension accuracy as the dependent variable. The model was significant: $F(3,76) = 5.408$, $p = .005$; adjusted R-squared = .143. Using a Bonferroni adjusted alpha level of .008 to give a family alpha level of .05, we then undertook six follow-up pairwise comparisons of means. Only one was statistically significant ($p = .001$)–namely, participants who received the entity-only ER model performed less well than participants who received the ontologically sound ER model.

Table 2. Means and Standard deviations for comprehension, problem-solving, and discrepancy-checking performance measures

	Comprehension		Problem Solving		Discrepancy Checking	
	Accuracy	Time	Accuracy	Time	Accuracy	Time
Ontologically Sound ER Model	74.50 (16.38)	6.70 (2.89)	59.75 (16.97)	29.27 (13.87)	50.63 (21.55)	16.20 (5.32)
Partially Ontologically Sound ER Model	62.50 (15.52)	8.11 (3.53)	54.5 (13.85)	28.05 (8.17)	37.97 (19.77)	12.97 (38.33)
Normalized ER Model	66.00 (19.84)	12.88 (5.24)	51.25 (17.24)	36.40 (13.10)	52.5 (21.28)	16.66 (6.46)
Entity-Only ER Model	50.50 (23.72)	11.40 (4.71)	49.25 (16.00)	31.92 (11.46)	43.91 (20.99)	16.23 (5.52)

Table 3. Pearson correlations among comprehension, problem-solving, and discrepancy-checking performance measures

	Comprehension Time	Problem-Solving Accuracy	Problem-Solving Time	Discrepancy-Checking Accuracy	Discrepancy-Checking Time
Comprehension Accuracy	-.177 (.116)	.407 (.000)	.266 (.017)	.338 (.002)	.006 (.955)
Comprehension Time		-.137 (.226)	.397 (.000)	.070 (.537)	.324 (.003)
Problem-Solving Accuracy			.176 (.119)	.452 (.000)	.124 (.273)
Problem-Solving Time				.217 (.053)	.435 (.000)
Discrepancy-Checking Accuracy					.297 (.007)

Next, we undertook an analysis of covariance (ANCOVA) with type of model as the factor, problem-solving accuracy as the dependent variable, and comprehension accuracy as the covariate. While comprehension accuracy was statistically significant as a covariate ($p = .001$), type of model was not statistically significant.

We then undertook another ANCOVA with type of model as the factor, discrepancy-checking accuracy as the dependent variable, and comprehension accuracy and problem-solving accuracy as the covariates. On this occasion, problem-solving accuracy was a statistically significant covariate ($p < .001$), but neither comprehension accuracy nor type of model was statistically significant.

For the three *time* measures, the MANOVA was also significant: $F(9, 180.247) = 3.528$, $p < .001$; Wilks' Lambda = .674, partial eta squared = .123. In this light, we again used the Roy-Bargmann stepdown analysis procedure to determine which of the three time measures were statistically significant. Once more, we entered the dependent variables into the step-down analysis following the order in which they had been measured in the experiment (i.e., comprehension time, problem-solving time, and discrepancy-checking time).

We first undertook a univariate analysis of variance (ANOVA) with type of model as the factor and comprehension time as the dependent variable. The model was significant: $F(3,76) = 9.267$, $p < .001$; adjusted R-squared = .239. Using a Bonferroni adjusted alpha level of .008 to give a family alpha level of .05, we then undertook six follow-up pairwise comparisons of means. Participants who received the ontologically sound ER model outperformed those who received the entity-only ER model ($p = .004$) and normalized ER model ($p < .001$). Furthermore, those who received the partially ontologically sound ER model outperformed those who received the normalized ER model ($p = .003$).

Next, we undertook an analysis of covariance (ANCOVA) with type of model as the factor, problem-solving time as the dependent variable, and comprehension time as the covariate. While comprehension time was statistically significant as a covariate ($p = .003$), type of model was not statistically significant.

We then undertook another ANCOVA with type of model as the factor, discrepancy-checking time as the dependent variable, and comprehension time and problem-solving time as the covariates. On this occasion, problem-solving time was a statistically significant covariate ($p < .002$), but neither comprehension time nor type of model was statistically significant.

3. Significant Findings

Our results show that the type of model had a significant effect for the comprehension task but not for the problem-solving or discrepancy-checking tasks. In terms of comprehension accuracy (interpretational fidelity), we found that participants who received the ontologically sound ER model outperformed those who received the entity-only ER model ($p = .001$). For comprehension time taken (interpretational efficiency), we found that participants who received the ontologically sound ER model outperformed those who received the normalized ER model ($p < .001$) and entity-only ER model ($p = .004$). Moreover, those who received the partially ontologically sound ER model outperformed those who received the normalized ER model ($p = .003$).

Although the comprehension *accuracy* results are somewhat muted in terms of support for our proposition, the *time taken* results are strongly supportive. We further explore the reasons why these results were found in the next section.

4. COGNITIVE PROCESS TRACING STUDY

We conducted a process tracing study to better understand the cognitive behaviour patterns of users of conceptual models and to help explain the outcomes we obtained in our laboratory experiment. We focus in particular on the ontological sound and normalized ER models to enable a comparison of the model motivated by our theory with the model most widely used in practice.

4.1 Design and Measures

We collected data about the cognitive processes of individuals who participated in our study using a verbal protocol technique. This technique requires individuals to verbalize their thoughts as they undertake some task (Ericsson & Simon, 1984). Cognitive process tracing is a recognized data gathering technique in cognitive psychology and information systems research. It is based on the assumption that humans consciously construct a representation of a problem and their detailed problem-solving strategies when they solve a problem. It also assumes that humans are able to access these strategies and verbalize them.

In this study we use the concurrent verbal protocol approach. Participants are asked to think aloud during the course of the task, thereby providing the researchers with direct access to their thought processes (Newell & Simon, 1972; Ericsson & Simon, 1984). Using verbal protocols provides a means to trace cognitive processes step by step, instead of relying on information about task outcomes or querying participants retrospectively about their cognitive processes. Our focus was on (a) understanding the cognitive behaviour of participants for those experimental tasks in which a significant difference was obtained, and (b) explaining why these outcomes occurred.

4.2 Materials

Four sets of materials were used in the study. The first was a "personal-profile" questionnaire to obtain information about participants' backgrounds. The second, third, and fourth sets of materials had been used in our prior laboratory experiment. The second comprised a summary of the ER symbols used in the diagrams provided to participants in the study. The third comprised two ER diagrams of alternative conceptual models of a sales order domain (one that is understood widely). We used only the ontological sound and normalized models from the laboratory experiment because the first model is motivated by our theory and the latter is the most widely used in practice. The fourth comprised five comprehension questions and five problem-solving questions.

4.3 Participants

Twelve participants took part in the study. All had at least three years' industry experience. They were selected on the basis that they would act as surrogate end users. They did not play an information technology role in their organisation, nor did they have previous data modelling experience.

4.4 Procedures

The materials were first pilot tested with two individuals who were not participants in the primary study. No concerns were identified. The primary study then commenced.

Participants in the primary study were first assigned randomly to one of the two alternative representation groups. Within each group, the sequence of tasks was altered for every second participant (comprehension followed by problem solving or problem solving followed by comprehension). Participants were then run singly through the task. When they arrived to undertake the task, they were asked to complete a consent form and the personal-profile instrument. The "speak-aloud" approach to data collection was then explained. A camcorder mounted on a tripod was focused on the ER models and used to (a) videotape participants as they indicated navigation of the models with a pencil, and (b) record participants' verbalizations.

Next, participants were given the document that explained the ER modelling symbols. They were permitted to discuss the symbols with the researchers until they indicated they felt confident with their meaning. Throughout the study, participants retained and could refer to the summary of the ER modelling symbols.

When participants indicated they were ready to begin, they were given either the "ontologically sound" or "normalized" ER diagram. They were then asked to work through the first task (either comprehension or problem solving). If periods of silence occurred, they were prompted to "speak aloud" to explain their cognitive behaviour. After a brief pause at the conclusion of the first task, participants were asked to work through the second task. At the conclusion of this task, participants were thanked and dismissed.

4.5 Coding Scheme

A coding scheme was established using the problem-solving literature (e.g., Newell & Simon 1972) and similar previous studies of data modelling (e.g., Batra & Davis, 1992; Chaiyasut, 1994; Shanks et al., 2008). This coding scheme comprised five cognitive behavior categories:

- *Understanding Question*: Includes reading the question, seeking clarification, identifying assumptions and constraints, and recognizing the problem posed.
- *Identifying Model Segment*: Includes locating appropriate parts of the model and matching them against key concepts in the question.
- *Articulating Model Semantics*: Includes verifying semantics of symbols in the model and re-reading the symbol summary.
- *Preparing Solution*: Includes developing solutions and simulating and revising solutions against the question.
- *Evaluation*: Includes selection of alternative answers and developing justifications.

4.6 Analysis of Protocol Data

All utterances on the videotapes were transcribed and partitioned into segments based on similar content. Video data was used to help identify start and end times for each segment. Each segment was then assigned to a cognitive behavior category within the coding scheme. Data was coded independently by two of the authors. Differences were reconciled.

Protocol data was further analyzed in three ways. First, the average time that participants spent in each of the five cognitive behavior categories was compared. This comparison indicates in which category the main differences occurred. Second, the proportion of time spent in each cognitive behavior category for each of ten equal time segments was compared. This comparison indicates which categories were prominent at different stages of the comprehension task. Third, the total number of transitions between each of the five categories was compared. This comparison indicates patterns in the sequence of cognitive behavior categories.

Figure 6 shows the average time that participants spent in each cognitive behavior category. Participants who received the ontologically sound model on average took 6.61 minutes to complete all five comprehension questions. Those who received the normalized ER model on average took 7.68 minutes to complete all five comprehension questions. The average time spent in identifying appropriate parts of the model and articulating model semantics was considerably less for participants who received the ontologically sound ER model. For example, participants who received the ontologically sound ER model spent 19.69 percent (compared to 31.30 percent for participants who received the normalized ER model) of their time in identifying appropriate parts of the model. This outcome suggests that participants found the normalized ER model more difficult to read and navigate. Furthermore, participants who received the ontologically sound ER model spent more of their time in understanding the model and evaluating their solutions, which suggests that they obtained a better comprehension of the model semantics.

Figure 7 shows the proportion of time spent in each cognitive behavior category for each of the ten time segments. The participants who received the ontologically sound ER model were able to identify appropriate parts of the model much earlier in the overall process. For example, in time segments 2 and 3, they spent about 50 percent of their time in the "identify" category, after which the proportion of their time in this category reduced sharply. In contrast, those participants who received the normalized ER model spent about 50 percent of their time in the "identify" category in time segments 3, 4, and 5, after which the proportion of their time in this category reduced at a slower rate. Similarly, preparing and evaluating solutions occurred much earlier for those participants who received the ontologically sound ER model.

Figure 8 shows the sequential dependencies between the five behaviour categories. The numbers above the dependency arrows are the total number of transitions between two categories. The thickness of the arrows indicates the intensity of the dependency. Overall, the most-common sequence for participants regardless of the type of model they received was to understand the question and to either identify the area in the model or directly prepare the

Figure 6. Average time spent in each behaviour category

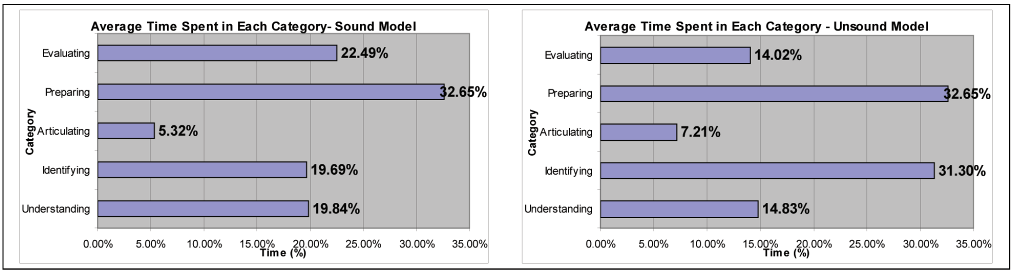


Figure 7. Proportion of time spent in each behaviour category

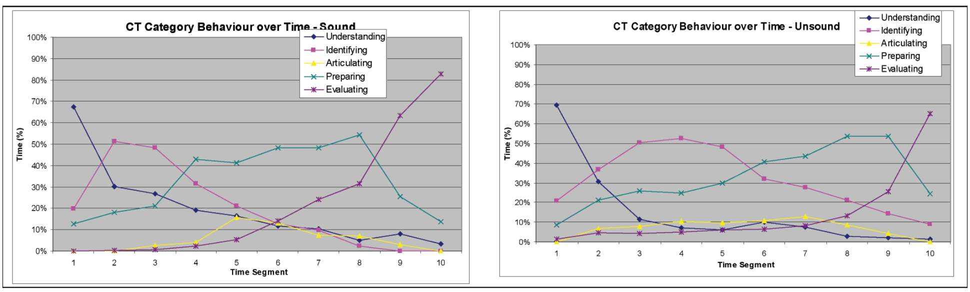
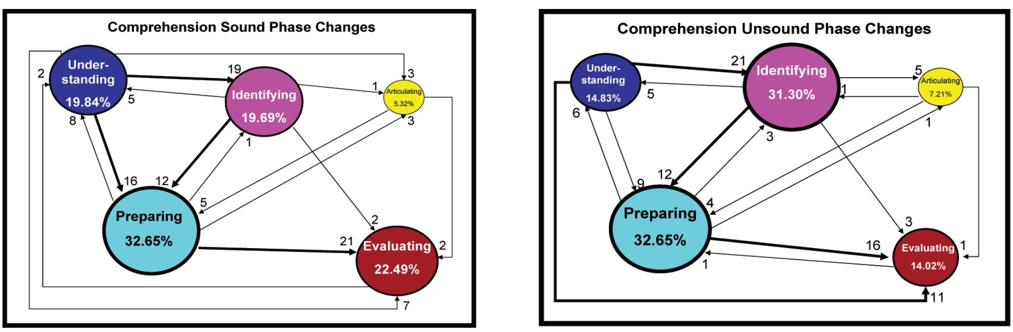


Figure 8. Total number of transitions between each behaviour category



solution before their final evaluation of the answer. Participants who received the ontologically sound model had less transition activity for the identifying model segment cognitive behavior category and more transition activity for the preparing solution cognitive behavior category. For example, they had 20 transitions in and 20 transitions out of the identifying model segment cognitive behavior category compared with 25 transitions in and 25 transitions out for participants who received the normalized ER model. They also had 43 transitions in and 32 transitions out of the preparing solutions model segment cognitive behavior category compared with 26 transitions in and 26 transitions out for participants who received the normalized ER model. This outcome is consistent with participants who received the ontologically sound model focusing more on solution preparation and evaluation rather than identifying appropriate parts of the model.

5. IMPLICATIONS OF THE RESEARCH

Our results provide some support for our proposition that classes of things and properties in general should be modelled explicitly as entity types and attributes. In this light, we argue that practitioners should be cautious when modelling properties in general as entity types and not attribute types. By failing to distinguish between classes of things and properties in general in the conceptual models they construct, they risk undermining users' understanding of the real-world phenomena being represented in the models. This understanding may be important to successfully accomplishing certain types of tasks that users of the models have to perform.

Our results also suggest that practitioners should be cautious about using the same type of model for both conceptual modelling and database-design purposes[9]. In this regard, relative to the normalized ER model, we have some evidence to indicate that the ontologically sound ER model better facilitates users' understanding of a domain. Given the way in which the normalized ER model has been constructed, however, we expect it is more suitable than the ontologically sound ER model as a means of supporting logical database design tasks. Moreover, because we could fairly easily transform an existing normalized ER model into an ontologically sound ER model, we expect that both types of model can co-exist satisfactorily. We predict that the ontologically sound ER model is best employed with users during requirements modelling and validation or comparisons of alternative models embedded within, say, different enterprise systems packages. On the other hand, we predict that the equivalent normalized ER model is best employed with database designers during implementation work carried out at later stages in the system-development process[10].

From a research perspective, our results strengthen a growing body of empirical work that supports the usefulness of ontological theories, particularly Bunge's (1977) ontological theory, as a means of predicting the strengths and weaknesses of alternative conceptual modelling methods (e.g., Weber, 1996; Green and Rosemann, 1996; Gemino, 1999; Opdahl & Henderson-Sellers, 2001; Parsons, 1996; Parsons & Wand, 2000; Bodart et al., 2001; Burton-Jones & Meso, 2006; Shanks et al., 2008). In the past, researchers have compared alternative conceptual modelling methods via omnibus feature comparisons or case studies (e.g., Olle et al., 1983). The results they have obtained using these approaches have been equivocal, which has motivated some researchers to call for better theory (e.g., Floyd, 1986). We argue that ontological theories can be used to address the shortcomings of these approaches because they allow us to pinpoint the strengths and weaknesses of alternative conceptual modelling methods. These predictions can then be tested using empirical research.

Our research also highlights the importance of the methods we employ to measure users' understanding of the phenomena represented by a conceptual model. Like prior research (e.g., Bodart et al., 2001), we found that some measures detected differences in the understanding obtained by users who studied a conceptual model. Other measures, however, detected no differences in user understanding. Presumably, users are eliciting different types of understanding when they study conceptual models. Unless appropriate measures are used, therefore, any differences in understanding that arise might not be detected.

6. LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Like most experimental and cognitive process tracing studies, our two studies are somewhat limited in scope and somewhat artificial. Future research might use alternative research methods, such as case studies and action research, to test our proposition in more-realistic settings (Siau & Rossi, 2007).

Our results are also limited by the validity and reliability of our measures. The fundamental construct that underlies our research, human understanding of a domain, is complex and multifaceted. We need better insights into the different types of understanding (e.g., surface-level versus deep-level) that users of conceptual models obtain when they employ a model. We also need better insights into how these different types of understanding support various tasks that users of conceptual models must undertake (e.g., responding to queries about a domain versus solving problems in the domain). Without these types of insights, our measures of user understanding will remain problematic.

Our results also suggest that humans might attend to different generic features of the real world, depending upon the task they must undertake. For instance, Shanks et al. (2008) found that users of a conceptual model attend to how composites and components are represented in the model when they have to solve problems about the domain represented by the model. On the other hand, we found that users of a conceptual model do *not* appear to be attending to whether a distinction is being made between classes of things and properties in general when they have to solve problems about the domain represented by the model. Instead, we found the distinction is important in relation to comprehension tasks that they have to perform. Moody (2002) and Weber (2003) have also argued the distinction is important if users have to undertake decomposition tasks. Thus, future research might examine what sorts of tasks, if any, are best supported by distinguishing between classes of things and properties in general in conceptual models.

ACKNOWLEDGMENT

Earlier and partial reports on the research described in this paper have been published in Shanks et al. (2003a), Shanks et al. (2003b), and Shanks et al. (2004). An Australian Research Council Discovery Grant funded this research. We thank Graeme Simsion for helpful comments on our research.

REFERENCES

Andersen, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261–295. doi:10.1016/S0022-5371(83)90201-3

Anderson, J. R., & Pirolli, P. L. (1984). Spread of activation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *10*(4), 791–798. doi:10.1037/0278-7393.10.4.791

Bacharach, S. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review*, *14*(4), 496–515. doi:10.2307/258555

Baddeley, A. (1994). The magical number seven: Still magic after all these years. *Psychological Review*, *101*(2), 353–356. doi:10.1037/0033-295X.101.2.353

Bodart, F., Sim, M., Patel, A., & Weber, R. (2001). Should optional properties be used in conceptual modelling? A theory and three empirical tests. *Information Systems Research*, *12*(4), 384–405. doi:10.1287/isre.12.4.384.9702

Boehm, B. (1981). *Software Engineering Economics*. Upper Saddle River, NJ: Prentice-Hall.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology*, *49*(October), 229–240.

Bunge, M. (1977). Treatise on Basic Philosophy: *Vol. 3. Ontology I: The Furniture of the World*. Boston: Reidel.

Burton-Jones, A., & Meso, A. P. (2006). Conceptualizing systems for understanding: An empirical test of decomposition principles in object-oriented analysis. *Information Systems Research*, *17*(1), 38–60. doi:10.1287/isre.1050.0079

Burton-Jones, A., Wand, Y., & Weber, R. (2009). Guidelines for Empirical Evaluations of Conceptual Modeling Grammars . *Journal of AIS*, *10*(6), 495–532.

Chen, P. P. S. (1976). The entity-relationship model: Toward a unified view of data. *ACM Transactions on Database Systems*, *1*(1), 9–36. doi:10.1145/320434.320440

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240–247. doi:10.1016/S0022-5371(69)80069-1

Connolly, T., & Begg, C. (2005). *Database Systems: A Practical Approach to Design, Implementation, and Management* (4th ed.). Harlow, UK: Addison-Wesley.

Date, C. J. (2003). *An Introduction to Database Systems* (8th ed.). Reading, MA: Addison-Wesley.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart.*, *13*, 319–340. doi:10.2307/249008

Denkel, A. (1996). *Object and Property*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511554575

Evermann, J., & Wand, Y. (2006). Ontological Modelling Rules for UML: An Empirical Assessment. *Journal of Computer Information Systems*, *46*(5), 14–19.

Floyd, C. (1986). A comparative evaluation of system development methods . In Olle, T. W., Sol, H. G., & Verrijn-Stuart, A. A. (Eds.), *Information System Design Methodologies: Improving the Practice* (pp. 19–54). Amsterdam, The Netherlands: North-Holland.

Gemino, A. (1999). *Empirical methods for comparing system analysis modelling techniques.* Unpublished doctorial dissertation, University of British Columbia.

Gemino, A., & Wand, Y. (2004). A framework for empirical evaluation of conceptual modelling techniques. *Requirements Engineering*, *9*, 248–260. doi:10.1007/s00766-004-0204-6

Green, P., & Rosemann, M. (2000). Integrated process modelling: An ontological evaluation. *Information Systems*, *25*(2), 73–87. doi:10.1016/S0306-4379(00)00010-7

Halpin, T. A. (2008). *Information Modeling and Relational Databases* (2nd ed.). San Francisco: Morgan Kaufman.

Hirschheim, R., Klein, H., & Lyytinen, K. (1995). *Information Systems Development and Data Modeling: Conceptual Foundations and Philosophical Foundations*. Cambridge, UK: Cambridge University Press.

Kent, W. (1978). *Data and Reality*. Amsterdam, The Netherlands: North-Holland.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, *59*, 43–64.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715–726. doi:10.1037/0022-0663.82.4.715

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review*, *63*, 81–97. doi:10.1037/h0043158

Moody, D. (2002). *Dealing with complexity: A practical method for representing large entity-relationship models*. Unpublished doctoral dissertation, University of Melbourne.

Moody, D. L., & Shanks, G. (1998). Improving the quality of entity-relationship models: An action research programme. *The Australian Computer Journal*, *30*, 129–138.

Nijssen, G. M. (1976). *Modelling in Database Management Systems*. Amsterdam: North Holland.

Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Olle, T. W., Sol, H. G., & Tully, C. J. (Eds.). (1983). *Information System Design Methodologies: A Feature Analysis*. Amsterdam, The Netherlands: North-Holland.

Parsons, J. (1996). An information model based on classification theory. *Management Science*, *42*(10), 1437–1453. doi:10.1287/mnsc.42.10.1437

Parsons, J., & Cole, L. (2005). What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modelling techniques. *Data & Knowledge Engineering*, *55*, 327–342. doi:10.1016/j.datak.2004.12.008

Parsons, J., & Wand, Y. (2000). Emancipating instances from the tyranny of classes in information modelling. *ACM Transactions on Database Systems*, *25*(2), 228–268. doi:10.1145/357775.357778

Parsons, J., & Wand, Y. (2008a). Using Cognitive Principles to Guide Classification in Information Systems Modeling. *Management Information Systems Quarterly*, *32*(4), 839–868.

Parsons, J., & Wand, Y. (2008b). A question of class. *Nature*, *455*, 1040–1041. doi:10.1038/4551040a

Popper, K. R. (1961). *The Logic of Scientific Discovery*. New York: Science Editions.

Rosemann, M., Davies, I., & Green, P. (2003). The very model of modern BPM. *Information Age, February/March*, 24-29.

Shanks, G., Nuredini, J., Tobin, D., Moody, D., & Weber, R. (2003a, June). Representing Things and Properties in Conceptual Modelling: an Empirical Evaluation. In *Proceedings of the European Conference on Information Systems,* Naples, Italy.

Shanks, G., Nuredini, J., Tobin, D., & Weber, R. (2003b). Representing Things and Properties in Conceptual Modeling: Understanding the Impact of Task Type. In *Proceedings of the Twenty-Third International Conference on Information Systems,* Barcelona, Spain (pp. 909-913).

Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2008). Representing part-whole relationships in conceptual modelling: An empirical evaluation. *Management Information Systems Quarterly*, *32*(3), 553–573.

Shanks, G., Weber, R., & Nuredini, J. (2004). Evaluating Conceptual Modelling Practices . In Rosemann, M., & Green, P. (Eds.), *Business Systems Analysis with Ontologies* (pp. 28–55). Hershey, PA: Idea Group Publishing.

Sia, S. K., & Soh, C. (2002). Severity assessment of ERP-organizational misalignment: Honing in on ontological structure and context specificity. In *Proceedings of the Twenty-Third International Conference on Information Systems*, Barcelona, Spain (pp. 723-729).

Siau, K. (2004). Informational and computational equivalence in comparing conceptual modeling methods . *Journal of Database Management*, *15*(1), 73–86.

Siau, K., & Rossi, M. (2007). *Evaluation techniques for systems analysis and design modelling methods – A review and comparative analysis*. Information Systems Journal.

Simsion, G., & Witt, G. (2001). *Data Modelling Essentials: Analysis* (2nd ed.). Design and Innovation.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Allyn and Bacon.

Teorey, T. J., Yang, D., & Fry, J. P. (1986). A logical design methodology for relational databases using the extended entity-relationship model. *ACM Computing Surveys*, *18*(2), 197–222. doi:10.1145/7474.7475

Wand, Y., Storey, V., & Weber, R. (1999). An Ontological Analysis of the Relationship Construct in Conceptual Modelling. *ACM Transactions on Database Systems*, *24*(4), 494–528. doi:10.1145/331983.331989

Wand, Y., & Weber, R. (1993). On the ontological expressiveness of information systems analysis and design grammars. *Journal of Information Systems*, *3*, 217–237. doi:10.1111/j.1365-2575.1993.tb00127.x

Wand, Y., & Weber, R. (2002). Information systems and conceptual modelling: A research agenda. *Information Systems Research*, *13*, 363–376. doi:10.1287/isre.13.4.363.69

Weber, R. (1996). Are attributes entities? A study of database designers' memory structures. *Information Systems Research*, *7*, 137–162. doi:10.1287/isre.7.2.137

Weber, R. (1997). *Ontological Foundations of Information Systems*. Melbourne, Australia: Coopers & Lybrand.

ENDNOTES

1 Note that we refer to conceptual (domain) modelling and not logical (e.g., relational database design) modelling.

2 Note that the "ontologically sound" model is not completely compliant with the principles of Bunge's (1977) ontological theory. In this study, we focus primarily on the distinction between classes of things and properties in general, and we are most concerned with representing intrinsic and mutual properties in general. We decided to retain the "Sales Order" and "Delivery" events in the "ontologically sound" model because the representation of events is a separate issue. Their deletion from the conceptual models might have confounded the results of our experiment.

3 Note that the conceptual models increase in complexity from the ontologically sound model (relatively simple) to the entity-only model (relatively complex). Although the

increasing complexity may make the models more difficult to read and interpret, it is a direct consequence of distinguishing things and properties in the models.

4 In terms of the notion of "physical independence," see Denkel (1996, pp. 34-35) on ways that things can be distinguished from properties.

5 We recognize that some changes of properties may result in a change in the "natural kind" of a thing (Bunge, 1977, p. 221).

6 Note, however, that the participants did not work in groups. Rather, they worked individually in the same room.

7 It proved impossible to record accurately the time taken for each discrepancy that a participant identified because, for example, participants vacillated back and forth between discrepancies as they attempted to find and articulate them. We ceased timing participants when they indicated they were done with the discrepancy-checking task.

8 We fully understand that this approach may lead to biases. The benefit, however, is that our scores are based on an in-depth understanding of our notes and our understanding of the participants' reactions as they undertook the experiment. Moreover, some of our scores are objective.

9 We use the term conceptual model to mean a representation of the data in an information system that is suitable for human understanding and is independent of any particular data management technology. This may be contrasted with a logical data model that represents that data in terms of a particular data management technology (e.g., relational) and a physical data model that takes into account specific storage structures and indexing mechanisms.

10 Note that these predictions are not based on the results of the experiment reported in this paper. They reflect our understanding of the consequences of the experimental results.

APPENDIX 1: EXPERIMENTAL TASKS

Example Comprehension Questions

1. Can an employee be assigned to manage more than one customer at a time?
   Yes / No / Not Sure
2. Can a customer belong to many postal area codes?
   Yes / No / Not Sure
3. Can an address have more than one region?
   Yes / No / Not Sure

Example Problem Solving Questions

1. The area code for all overseas telephone numbers has changed. Can we always identify the customer contact numbers for the customers located overseas?

    1. Possible / Not possible / Not Sure

    Explanation:

    2. An Ontological Plastics supplier wishes to send samples of new and improved hoses to customers who regularly order hoses. Can we determine the number of hoses each customer has had delivered in the previous 3 months and the date of each delivery?

    1. Possible / Not possible / Not Sure

    Explanation:

    3. A customer was delivered industrial piping, which has developed cracks in it 3 days after delivery. The customer wants a refund and the faulty piping collected by Ontological Plastics. Does the model allow an empty delivery truck to be sent to collect the goods?

    Possible / Not possible / Not Sure

    Explanation:

Example Paragraphs from Discrepancy Checking Task

Modeller: *Thank you for meeting with me. I was hoping that each of you could identify the requirements of your business area for me so I can start developing a model for you.*

User: *Sure, happy to help. I'll start, because I work in the customer side of the business. We deal with a number of customers, ranging from companies to individuals. Because of our large customer base, we feel it is important for each customer to have a specific relationship with an employee, whose role is to ensure customer satisfaction.*

All our customers have an employee assigned to manage them at a particular point in time. With large customers, multiple employees may be assigned to them.

As part of our customer satisfaction focus, it is imperative that we keep an accurate record of customer details, especially: credit limits, contact numbers, Australian Business Numbers, industry, date of registration etc. Also, when placing sales orders customers sometimes have delivery instructions, which we pass to the delivery team.

Graeme Shanks is Professor in the Department of Information Systems at the University of Melbourne. His research interests include conceptual modeling, data quality and the impact of business analytics systems on firm performance. He was recently awarded an Australian Professorial Fellowship from the Australian Research Council.

Daniel Moody is a consultant in data management. He has held several academic positions at universities throughout Australia and Europe. His research interests include conceptual modeling, software quality, data warehousing and medical informatics.

Daniel Tobin is a marketing manager at the University of Melbourne. He was previously a data management consultant and a research assistant at the University of Melbourne.

Jasmina Nuredini is a consultant in data management. She was previously a research assistant at the University of Melbourne.

Ron Weber is Dean of the Faculty of Information Technology, Monash University. His primary research interests lie in the ontological foundations of information systems. He is a past president of the Association for Information Systems and a past editor-in-chief of the MIS Quarterly. *In 2000, he won the Australian Prime Minister's Award for University Teacher of the Year.*